7-2015

Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development

Hongjun Liu
Sichuan Agricultural University

Xuerong Yang
Sichuan Agricultural University

Xinhui Liao
BGI-Shenzhen

Tao Zuo
Iowa State University

Cheng Qin
Sichuan Agricultural University

See next page for additional authors

Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development

Abstract

Background: Development of the maize (Zea mays L.) female inflorescence (ear) has an important impact on corn yield. However, the molecular mechanisms underlying maize ear development are poorly understood.

Results: We profiled and analyzed gene expression of the maize ear at four developmental stages: elongation phase (I), spikelet differentiation phase (II), floret primordium differentiation phase (III), and floret organ differentiation phase (IV). Based on genome-wide profile analysis, we detected differential mRNA of maize genes. Among the ~6,800 differentially expressed genes (DEGs), 3,325 genes were differentially expressed in stage II, 3,765 genes in III, and 1,698 genes in IV, compared to its previous adjacent stages, respectively. Furthermore, some of DEGs were predicted to be potential candidates in maize ear development, such as AGAMOUS (GRMZM2G052890) and ATFP3 (GRMZM2G155281). Meanwhile, some genes were well-known annotated to the mutants during maize inflorescence development such as compact plant2 (ct2), zea AGAMOUS homolog1 (zag1), bearded ear (bde), and silky1 (si1). Some DEGs were predicted targets of microRNAs such as microRNA156. K-means clustering revealed that the DEGs showed 18 major expression patterns. Thirteen transcriptional factors from 10 families were differentially expressed across three comparisons of adjacent stages (II vs. I, III vs. II, IV vs. III). Antisense transcripts were widespread during all four stages, and might play important roles in maize ear development. Finally, we randomly selected 32 DEGs to validate their expression patterns using quantitative reverse-transcription polymerase chain reaction (qRT-PCR). The results were consistent with those from Solexa sequencing.

Conclusions: DEGs technique had shown an advantage in detecting candidates, and some transcription factors during maize ear development. RT-PCR data were consistent with our sequencing data and supplied additional information on ear developmental processes. These results provide a molecular foundation for future research on maize ear development.

Keywords

maize, DGE, ear development, microRNA156, TFs

Disciplines

Agricultural Science | Agronomy and Crop Sciences | Genomics | Molecular Genetics | Plant Breeding and Genetics

Comments

This is a manuscript of an article published as Liu, Hongjun, Xuerong Yang, Xinhui Liao, Tao Zuo, Cheng Qin, Shiliang Cao, Ling Dong et al. "Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development." *Genomics* 106, no. 1 (2015): 52-60. doi: 10.1016/j.ygeno.2015.03.005. Posted with permission.

Authors

Hongjun Liu, Xuerong Yang, Xinhui Liao, Tao Zuo, Cheng Qin, Shiliang Cao, Ling Dong, Huangkai Zhou, Yongzhong Zhang, Sisi Liu, Yaou Shen, Haijian Lin, Thomas Lubberstedt, Zhiming Zhang, and Guangtang Pan

Genome-wide comparative analysis of digital gene expression tag profiles during maize ear development

Hongjun Liu[1†], Xuerong Yang[2†], Xinhui liao[3†], Tao Zuo[4], Cheng Qin[1,5], Shiliang Cao[6], Ling Dong[7], Huangkai Zhou[3], Yongzhong Zhang[1], Sisi Liu[1], Yaou Shen[1], Haijian Lin[1], Thomas Lübberstedt[8], Zhiming Zhang[1*], Guangtang Pan[1*]

Affiliations:

[1]Maize Research Institute of Sichuan Agricultural University / Key Laboratory of Biology and Genetic Improvement of Maize in Southwest Region, Ministry of Agriculture, Chengdu, 611130, China.

[2] Animal Nutrition Institute, Sichuan Agricultural University, Ya'an 625014, China

[3] BGI-Shenzhen, Shenzhen, 518083, China

[4] Interdepartmental genetics program, Iowa State University, Ames, 50011, USA

[5] Zunyi Academy of Agricultural Sciences, Zunyi, 563102, China.

[6] Maize Research Institute of Heilongjiang Academy of Agricultural Sciences, 150086, Harbin, China

[7] Department of Agronomy, Northeast Agricultural University, 150030, Harbin, China

[8] Department of Agronomy, Iowa State University, Ames 50011, USA

[*]E-mail: pangt@sicau.edu.cn or zzmmaize@gmail.com

[†]These authors contributed equally to this work.

Abstract

Background: Development of the maize (*Zea mays* L.) female inflorescence (ear) has an important impact on corn yield. However, the molecular mechanisms underlying maize ear development are poorly understood.

Results: We profiled and analyzed gene expression of the maize ear at four developmental stages: elongation phase (I), spikelet differentiation phase (II), floret primordium differentiation phase (III), and floret organ differentiation phase (IV). Based on genome-wide profile analysis, we detected differential mRNA of maize genes. Among the ~6,800 differentially expressed genes (DEGs), 3,325 genes were differentially expressed in stage II, 3,765 genes in III, and 1,698 genes in IV, compared to its previous adjacent stages, respectively. Furthermore, some of DEGs were predicted to be potential candidates in maize ear development, such as *AGAMOUS* (GRMZM2G052890) and *ATFP3* (GRMZM2G155281). Meanwhile, some genes were well-known annotated to the mutants during maize inflorescence development such as compact plant2 (*ct2)*, zea AGAMOUS homolog1 (*zag1)*, bearded ear (*bde)*, and silky1 (*si1)*. Some DEGs were predicted targets of microRNAs such as microRNA156. K-means clustering revealed that the DEGs showed 18 major expression patterns. Thirteen transcriptional factors from 10 families were differentially expressed across three comparisons of adjacent stages (II *vs.* I, III *vs.* II, IV *vs.* III). Antisense transcripts were widespread during all four stages, and might play important roles in maize ear development. Finally, we randomly selected 32 DEGs to validate their expression patterns using quantitative reverse-transcription polymerase chain reaction (qRT-PCR). The results were consistent with those from Solexa sequencing.

Conclusions: DEGs technique had shown an advantage in detecting candidates, and some transcription factors during maize ear development. RT-PCR data were consistent with our sequencing data and supplied additional information on ear developmental processes. These results provide a molecular foundation for future research on maize ear development.

Keywords: maize; DGE; ear development; microRNA156; TFs;

Introduction

Maize (*Zea mays* L.) is one of the most important crops and widely used as a model plant. Inflorescence and flower development are critically important for high yields in maize. Maize ears contain a low concentration of nitrogen, which can make good use of efficiency in aid sustainable production without adding more soil fertility after harvest [1]. Various mutants have been discovered, providing insights into molecular processes involved in ear development [2-7]. However, the understanding of the maize ear developmental dynamics at the transcriptome level is limited. Only a few studies have undertaken large-scale gene expression analysis of the maize ear, including (i) evaluation of sequence-based expression profiles during reproductive organ development [8], (ii) a study of the effect of water-deficit on immature maize ear development [9], and (iii) discovery of novel microRNAs during maize ear development [10].

The B73 sequence assembly [11] enables further analysis of maize ear development at a genome-wide transcriptome level. Owing to the dramatic decrease of sequencing costs and development of rapid and robust experimental procedures, sequencing can be used for cost-efficient high-throughput profiling analysis. For instance, by using digital gene

expression (DGE) [12-15] and RNA-Seq [16-18] analyses, new genes have been discovered [19]. Furthermore, these technologies are useful for estimating overall gene expression at different development stages or in different tissues [12,20] and in response to abiotic stress [21,22]. Considering the importance of ears in maize production, it is of great importance to understand molecular mechanisms involved in maize ear development in detail.

Thus, the objective of this study was to undertake a genome-wide comparative analysis of gene expression profiles to obtain an improved understanding of the molecular mechanisms of maize ear development during four developmental stages (the growth point elongation, spikelet differentiation, floret primordium differentiation, and floret organ differentiation phases) [23] using a DGE approach. Ears of maize from the four developmental stages were used to study the dynamics of mRNA expression. qRT-PCR was performed using randomly selected DEGs, in order to validate their expressions across different developmental stages. The K-means clustering method was employed to further determine the co-expression of genes involved in maize ear development.

Materials and Methods

Plant cultivation and sample collection

Seeds of the maize inbred line 18-599 (Maize Research Institute, Sichuan Agricultural University, Chengdu, China) were grown in a growth chamber at 24°C/18°C (day/night) with 12 h illumination per day. Ears were collected as described previously [10] at four developmental stages: the growth point elongation, spikelet differentiation, floret primordium differentiation, and the floret organ differentiation phases. In brief, ears were

manually collected at the four developmental stages. All the samples were harvested and immediately frozen in liquid nitrogen, and stored at -80°C until used for RNA isolation.

Digital expression library preparation and sequencing

Total RNA from the maize inflorescences at each developmental stage were isolated using TRIzol (Invitrogen, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. For RNA library construction and deep sequencing, equal quantities of RNA were pooled for each developmental stage. Approximately 6 μg of total RNA representing each library were sequenced using Illumina HiSeq™ 2000 System as described [15]. The DGE libraries were constructed using Illumina Gene Expression Sample Prep Kit according to the manufacturer's instructions. Briefly, mRNA were isolated and purified from total RNA using Oligo (dT) magnetic beads, and synthesized to be the first and second-strand cDNA. Then the bead-bound cDNA was subsequently digested with restriction enzyme *NlaIII*, and two adaptors were ligated to the 5' and 3' ends of the tags, respectively. After 15 cycles of linear PCR amplication, 95bp fragments were purified. Then the single-chain molecules were fixed onto the Illumina Sequencing Chip (flowcell) and sequenced with the method of sequencing by synthesis. Finally, millions of raw reads of each library were generated with sequencing length of 35 bp.

Sequencing data and Differentially expressed gene analysis

Raw data were filtered to remove the adaptors, low quality tags, and tags with one copy number. Then clean reads were used for further analysis. In brief, clean reads were aligned to the maize reference genome [11] (B73 RefGen_v3 (May 2012)) using SOAP2 software [24], allowing only 1bp mismatch. The unambiguous tags were annotated. Both sense and antisense sequences were used in the data analysis. To analyze the gene

expression, the number of clean reads for each library was counted and then normalized to tags (reads) per million. To detect DEGs, statistical analyses among libraries were performed following the formula as described [20], where false discovery rate (FDR) was used to determine the threshold of the P value in multiple test and statistical analysis [25]. Significance of differential gene expression was determined at a false discovery rate < 0.001 and an absolute value of log2-ratio ≥ 1.

Global analysis of differentially expressed genes

To further annotate and analyze the DEGs, a user-driven tool MapMan [26] was used to assign DEGs to functional categories of metabolic pathways or other processes, and the genes were grouped by developmental dynamics using the K-means clustering algorithm as described [27]. To further identify the significantly enriched metabolic pathways in DEGs, Gene Ontology and enrichment analyses with KEGG annotation were performed using the genes within each cluster, where the formula used in this study is as follows [20]:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

In this formula, N indicates the number of DEGs, n indicates the number of genes within each cluster in N. M indicates the number of the DEGs with specific GO/KEGG annotations, and m indicates the number of genes within each cluster in M.

Quantitative real-time PCR analysis

To validate the DEGs, quantitative real-time PCR was performed for 32 randomly selected DEGs, which were the same as those used for the DGE genome-wide comparative analysis. Real-time PCR was performed using the SYBR Premix Ex Taq™ protocol (TaKaRa Biotechnology, Dalian, China**)** on an Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). For each sample, measurements were performed in triplicate, with three biological replicates, and the average cycle thresholds (Ct) were used to determine fold-change. *18S rRNA* (forward primer: 5′-ATGTTCCGTGGCAAGATGAG-3′, reverse primer: 5′-CATTGTTGGGAATCCACTC-3′) was employed as an endogenous control. Primers were designed using the Primer Premier 5.0 (http://www.premierbiosoft.com/index.html) and Oligo 6 programs (http://www.oligo.net) (Table S8). Thermal cycle conditions were as follows: 2 min at 95°C followed by 40 cycles of 15 s at 95°C, 15 s at 56-57°C, and 15 s at 72 °C. Statistical analysis was conducted using the $2^{-\triangle\triangle Ct}$ method as described previously [28].

Data access

RNA sequencing data have been deposited at NCBI under the accession number GSE49805.

Results

Sequence identification and analysis

Library construction and sequence analysis were conducted as described previously [20]. Generally, around 16.8 million high-quality raw reads were generated. After performing quality-control measures, 16.1 million clean tags were obtained for all four

stages (Tables 1 and S1). Briefly, after removing low-quality and contaminating reads, clean tags were retained for further analysis. Subsequently, the 16.1 million clean tags were aligned against the maize genome (B73 RefGen_v3 (May 2012)). The percentage of clean tags in the raw data for each developmental stage was 93.62%, 96.81%, 96.54%, and 96.56%, respectively. About 69.86% clean tags were mapped to the B73 reference genome with an average of 76.10% genes covered. The occurrence of unmapped tags was probably due to incompleteness of the maize genome sequence data. Most tags were aligned to genic regions and the genic distribution of reads from mRNA reference sequences at the four developmental stages (I-IV) showed that the majority of tags (87.58%, 87.23%, 90.30%, and 90.95%, respectively) were mapped to exon regions and the remainder were distributed among introns, intergenic regions and repeat regions (Figure S1).

The majority of transcripts were expressed at all four stages (Figure 1A, Table S2). The number of sense (Figure 1A) and antisense (Figure 1B) transcripts overlapping at all four stages were 11,970 and 4416, with a cutoff for expression at each stage of one tag per million. The number of genes that showed both sense and antisense expression were 7230, 7052, 6918, and 6571 (Figure 1C) for each developmental stage, and 10,456 across all stages. Of those sense genes detected, only 74 genes were expressed uniquely in stage I, and this number was even lower than that of the other three developmental stages, which suggested that more genes were involved in maize ear development during the ear developmental transition.

Analysis of differentially expressed genes and validation by qRT-PCR

Generally, all tags that were mapped to genes (sense and anti-sense) were used for differential expression analysis combined with the DGE method for a genome-wide comparative analysis of data for the four developmental stages. Comparative gene expression analyses were used for estimation of gene expression levels at the four developmental stages (Table S3). We calculated the number of tags corresponding to each gene in each library to estimate gene expression levels and compare the difference in fold-change between the developmental stages [20]. Transcripts that showed differential expression levels are shown in Figure S3; the blue dots are defined as "no difference in expression" representing genes that are differed by less than 2-fold change between two libraries at a threshold of log2 ratio ≥ 1. The up-regulated (red dots) and down-regulated genes (green dots) indicate significantly DEGs (Figures 2A and S2, Table S4). In total, the number of genes differentially expressed between two stages was as follows: 3325 between stages I and II (36% up- and 64% down-regulated in stage II), 4735 between stages I and III (57% up- and 43% down-regulated in stages III), 6398 between stages I and IV (46% up- and 54% down-regulated in stage IV), 3765 between stages II and III (71% up- and 29% down-regulated in stage III), 5178 between stages II and IV (60% up- and 40% down-regulated in stage IV), and 1698 between stages III and IV (35% up- and 65% down-regulated in stage IV).

To better understand the dynamic changes of gene expression in maize ear development during the four stages, further analyses of the DEGs were performed, especially those genes up- or down-regulated gradually following ear development (II *vs.* I, III *vs.* II, and IV *vs.* III; Table S4). Among the DEGs identified, 1201, 2690, and 594 genes were up-regulated in stages II, III, and IV, respectively, compared with their own

preceding stage. In contrast, the number of down-regulated genes was 2124, 1075, and 1104 in stages II, III, and IV, respectively (Figure 2A, Table S4). This suggests that DEGs were less abundant in stages II and IV, whereas a higher number (uniquely 2217 DEGs in III *vs.* II, Figure 2B) of DEGs were involved in ear development during stage III (Figure 2B), indicating that more genes were involved in stage I (growth point elongation phase) and stage III (floret primordium differentiation phase) during maize ear development. Furthermore, the expression patterns of nine DEGs were illustrated in Figure S4. Interestingly, some genes were well-known annotated to the mutants during maize inflorescence development such as compact plant2 (*ct2)*, zea AGAMOUS homolog1 (*zag1)*, bearded ear (*bde)*, and silky1 (*si1)* [29].

To confirm the expression patterns determined by Solexa RNA-sequencing analysis, we used qRT-PCR analyses to analyze the expression of 32 randomly selected genes (Figure 4). Although the Solexa log2-fold values of the 32 genes showed slight variation compared with the corresponding values from the qRT-PCR analyses, the expression data from the Solexa RNA-Seq analysis were closely positively correlated (most Pearson correlation coefficients were higher than 0.8) with those obtained by qRT-PCR (Figures 3 and S3), which indicated that the deep sequencing and RT-PCR data were well consistent.

Gene ontology enrichment and global analysis of the gene expression profiles

On the basis of the DEGs and gene ontology (GO) annotation, GO enrichment analysis assigned the DEGs (with q-value < 0.001 significantly enriched, and |log2 ratio| ≥ 1) to one of three functional categories: cellular component, molecular function, and biological process. In general, approximately 37% DEGs were successfully assigned to at

least one GO term among the following stage comparisons: II *vs.* I, III *vs.* II, and IV *vs.* III. During stage II, 1208 (36.33%) out of 3325 DEGs were overrepresented in at least one GO term. During stage III, 1380 (36.65%) out of 3765 DEGs were overrepresented in least one GO term. In stage IV, only 644 (37.93%) out of 1698 genes were successfully assigned to at least one GO term.

We identified ~6800 out of 25,800 genes that were differentially expressed among developmental stages, representing 26% of the ear transcriptome. Using MapMan [26], we identified 18 clusters (K1 to K18; Figure 4A, Table S5). Most clusters showed significant enrichment for particular GO terms and KEGG pathways (Figure 4B, C). In general, clusters K1 and K7 displayed similar expression patterns across the four developmental stages, which were consistent with the results of the GO enrichment analyses. Most overrepresented GO terms were shared between clusters K1 and K7, such as chromatin organization, organelle organization, and cellular component organization or biogenesis (Figure 4B). In addition, ribosome pathway was also overrepresented in both clusters K1 and K7. Interestingly, the *carboxy-lyase* (GRMZM2G159149) in cluster K1 was found to be uniquely, significantly up-regulated in stage III, and significantly down-regulated in stages II and IV (Table S4), which indicated *carboxy-lyase* may have an important role in the floret primordium differentiation phase during maize ear development. However, only one GO term (response to stress) was significantly enriched in cluster K2.

Resolving transcription factors among differentially expressed genes

A primary objective was to identify genes that encode transcription factors (TFs) and resolve the dynamics of TF accumulation during ear development. We retrieved putative

orthologs of maize genes based on information from the EnsemblCompara gene trees [30] at Gramene (http://maizesequence.org), PlantGDB (http://plantgdb.org), and the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). We then queried known plant TFs in the Plant Transcription Factor Database (v2.0, http://planttfdb.cbi.edu.cn/) and identified 1,522 maize TFs with sequence similarities to known plant TFs that were expressed during at least one developmental stage (Table S6). In total, 242 TFs were differentially expressed between II *vs.* I, 345 between III *vs.* I, 445 between IV *vs.* I, 239 between III *vs.* II, 358 between IV *vs.* II, and 132 between IV *vs.* III (Table S7). In total, 13 TFs from 10 TF families (B3, bZIP, ERF, HD-ZIP, LBD, MIKC, MYB, NAC, SBP, and TCP) were differentially expressed across comparisons of adjacent stages during all four developmental stages of maize ear development (Table 3).

SBP-box genes, which encode a class of zinc finger-containing TFs, are important regulators with various functions during maize development [31]. In this study, *SBP-domain protein 5* (GRMZM2G160917, 1.72 log2 fold change for III *vs.* II, 1.85 log2 fold change for II *vs.* I, *q*-value < 0.001) and *SBP-domain protein 6* (GRMZM2G307588, 1.37 log2 fold change, *q*-value < 0.001) were significantly up-regulated in stages III & II, and stage IV compared to its previously adjacent stages, respectively (Figure 5, Table S7). Interestingly, both genes were predicted targets of microRNA156, which plays an important role in maize ear development [10] and tomato growth development [32]. The result further indicates the potential roles of microRNA156 and SBP in maize ear development. As a TF, SBP (GRMZM2G109354) displayed significantly differential expression patterns during different developmental stages (Figure 5C), which suggested SBP may have different regulatory roles in each stage of maize ear development.

Antisense transcripts detected for many genes

Antisense transcripts have been identified and predicted from maize [33]. Antisense transcription is a common phenomenon in maize (Figure 1B, C) and was widespread at all four stages. Interestingly, a large number of antisense transcripts were uniquely expressed in stage I (1232 genes), whereas substantially fewer (692) genes were expressed in the antisense direction at the other stages (Figure 1C). During stage I, 7230 genes were expressed in both sense and antisense direction (Figure 1C). The distribution pattern of antisense transcripts across the four developmental stages was similar to the overall transcript distribution (both antisense and sense, Figure 1C).

Annotation with GO terms was performed for the antisense transcripts detected at all four developmental stages (Table S2). As described above, more unique antisense transcripts were detected during stage I. Those unique antisense transcripts during stage I were assigned to 41 GO terms for cellular component such as mitochondrion (GO:0005739), 144 for molecular function such as oxidoreductase activity (GO:0016491), and 109 for biological process such as protein catabolic process (GO:0030163).

Discussion

In this study, ~6800 DEGs were identified across various comparisons of developmental stages during maize ear development. Previous studies suggested that the expression and function of the floral homeotic gene *AGAMOUS* (*AG*) was responsible for normal floral development and floral organ identity transformation [34,35]. In the present study, the homolog of the Arabidopsis gene *AGAMOUS* (GRMZM2G052890) (9.21 log2

fold change, *q*-value < 0.001) was the most significantly up-regulated and annotated gene in stage III compared with stage II, which suggests this gene may have an important role in ear development during stage III.

Among those DEGs validated by qRT-PCR, the gene GRMZM2G007025 displayed differential expression between stages III and IV. During stage III, GRMZM2G007025 was down-regulated (2.377 log2 fold change, *q*-value < 0.001) in contrast with stage II. It was up-regulated (2.382 log2 fold change, *q*-value < 0.001) during stage IV compared with stage III. Interestingly, the mRNA level of GRMZM2G007025 was higher in stage IV than that in stage III (Figure 3), which is consistent with our DGE results. This indicates that GRMZM2G007025 may be a candidate gene for better understanding the mechanism of maize ear development.

Metal ions such as zinc, copper, and iron are essential for plant growth. Dykema *et al*. (1999) [36] characterized ATFP3 (*Arabidopsis thaliana farnesylated protein 3*) as an important factor that binds to transition metal ions. In the current study, *ATFP3* was the most significant and annotated up-regulated gene (GRMZM2G155281, 2.72 log2 fold change, *q*-value < 0.001, Table S4) during stage II compared with stage I by ranking the significance, which indicated that additional metal ions might be required for maize ear development during stage II. In addition, soluble inorganic pyrophosphatase (*PPase*) (GRMZM2G104918) plays an important role in the adaptation of *Phaseolus vulgaris* to phosphate starvation [37]. The different expression patterns were found for *PPase* during developmental stages II and III. *PPase* was significantly down-regulated (1.27 log2 fold change, *q*-value < 0.001) in stage II in contrast with stage I, whereas it was up-regulated (3.18 log2 fold change, *q*-value < 0.001) at stage III *vs.* stage II. *PPase* might thus be a

candidate gene for prediction of the physiological signal of phosphorus during maize ear development.

In plants, microRNAs play important regulatory roles in many aspects of plant biology, including metabolism, growth, and stress response [10]. Liu *et al.* reported that zma-miRNA156 was differentially expressed during maize ear development [10]. In accordance with the previous study, 19 DEGs were predicted to be the targets of microRNA156 and microRNA319 (stage II *vs.* I), microRNA156, 160, 164, 167, 390, and 394 (stage III *vs.* II), and microRNA156, 160, 319, and 529 (stage IV *vs.* III) during the four developmental stages (Table 2).

To further investigate the co-expression pattern of DEGs involved in maize ear development, we identified several clusters that were uniquely significantly enriched in GO terms, such as cellular aromatic compound metabolic process and microtubule-based process, in clusters K5 and K13 (Figure 4B), respectively. These two clusters were also significantly involved in KEGG pathways, such as phenylalanine, tyrosine, and tryptophan biosynthesis of secondary metabolites for K5, and proteasome for K13 (Figure 4C). Sixteen genes were significantly differentially expressed in cluster K13 across the four developmental stages (Table 2), of which most of the genes showed a unique expression pattern (either up- or down-regulated) during stage III.

In plants, protein phosphorylation is an important regulatory mechanism. Histidine-containing phosphotransfer proteins (*HPts*, GRMZM2G016439) participate in hormone signal transduction in higher plants [38]. These genes were significantly down-regulated in stage III *vs.* II, but up-regulated during stage II *vs.* I and stage IV *vs.* III, which suggested that phosphorylation and hormone signal transduction were more active in

stages II and IV and *HPts* might be a potential marker of phosphorylation status during maize ear development. However, this suggestion needs to be further investigated. Moreover, maize *proteinase inhibitor* (GRMZM2G028656) in cluster K2 was significantly down-regulated across all of the three comparisons, especially during stage IV (8.79-fold in stage IV *vs.* III, 1.13-fold in stage III *vs.* II, and 1.47-fold in stage II *vs.* I). Plant proteinase inhibitors play an important role in the insect resistance of transgenic plants [39]. Maize *proteinase inhibitor* was significantly annotated in the serine-type endopeptidase inhibitor activity (GO:0004867), peptidase activity (GO:0008233), and response to wounding (GO:0009611).

To better explain the molecular mechanisms of maize ear development, 13 genes encoding transcriptional factors were identified across all four stages. The B3 DNA-binding domain, a plant-specific domain, is found throughout flowering plants [40]. The Arabidopsis B3 domain protein VERNALIZATION1 (VRN1) is involved in processes essential for development [40]. In the present study, B3 DNA binding domain-containing protein (GRMZM2G065496) significantly up-regulated in stage III (*vs.* II, 1.51-fold, *q*-value < 0.001), and down-regulated in stage II (*vs.* I, 1.08-fold, *q*-value < 0.001), and stage IV (*vs.* III, 1.00-fold, *q*-value < 0.001), which indicates B3 may have an important role in the development of floret primordium differentiation (stage III).

Antisense transcripts have been reported in various biological processes, including translation regulation and RNA interference [33]. In the present study, several additional processes were significantly identified from unique antisense transcripts during stage I, including oxidation reduction (GO:0055114), response to oxidative stress (GO:0006979), regulation of protein metabolic process (GO:0051246), phosphatidylcholine metabolic

process (GO:0046470), inflorescence development (GO:0010229), and microtubule bundle formation (GO:0001578). Thus, the antisense transcripts detected in this study may play important roles in maize ear development.

Conclusions

In this study we undertook a comprehensive transcriptome analysis and identification of DEGs during maize ear development using a Solexa sequencing platform. In total, we identified 3325 genes that were differentially expressed during the spikelet differentiation phase, 3765 genes during the floret primordium differentiation phase, and 1698 genes during the floret organ differentiation phase, compared to its previously adjacent stages., respectively. Some of these DEGs were predicted to play important roles in maize ear development, such as *AGAMOUS* (GRMZM2G052890*)* and *ATFP3* (GRMZM2G155281). Interestingly, some genes were well-known annotated to the mutants during maize inflorescence development such as compact plant2 (*ct2)*, zea AGAMOUS homolog1 (*zag1)*, bearded ear (*bde)*, and silky1 (*si1)*. In accordance with our previous study [10], several DEGs were predicted to be targets of microRNAs. MicroRNA156 appears to be a key microRNA in maize ear development. As predicted targets of microRNA156, SBP-box genes such as *SBP-5* and *SBP-6* were indicated to play differentially important roles in different stages of maize ear development. K-means clustering revealed 18 major expression patterns. From the analysis of TFs, we also identified 13 transcriptional factors from 10 TF families (B3, bZIP, ERF, HD-ZIP, LBD, MIKC, MYB, NAC, SBP, and TCP) that were differentially expressed across three adjacent comparisons (II *vs.* I, III *vs.* II, and IV *vs.* III) of four developmental stages of maize ear development. Antisense transcripts were widespread during all four stages, and

especially a large number of antisense transcripts existed at developmental stage I, thus the antisense transcripts detected in this study may play important roles in early stage of maize ear development. Understanding maize ear development is critical for improvement of maize production. Thus, identification and characterization of important genes and regulators at the four developmental stages will contribute to an improved understanding of the molecular mechanisms responsible for maize ear development.

Author Contributions

Conceived and designed the experiments: GP ZZ HL. Performed the experiments: HL CQ YZ SL. Analyzed the data: HL XY XL CQ TZ HZ TL. Contributed reagents/materials/analysis tools: HL GP ZZ YS HL. Wrote the paper: HL XY SC LD TL GP ZZ. All authors have read and approved the manuscript for publication.

Reference

1. Jansen C, Lübberstedt T (2012) Turning Maize Cobs into a Valuable Feedstock. BioEnergy Research 20-31.
2. Bommert P, Je BI, Goldshmidt A, Jackson D (2013) The maize Galpha gene COMPACT PLANT2 functions in CLAVATA signalling to control shoot meristem size. Nature 502: 555-558.
3. Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, et al. (2006) ramosa2 encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. Plant Cell 18: 574-585.

4. Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ (2002) The control of spikelet meristem identity by the branched silkless1 gene in maize. Science 298: 1238-1241.

5. Gallavotti A, Long JA, Stanfield S, Yang X, Jackson D, et al. (2010) The control of axillary meristem fate in the maize ramosa pathway. Development 137: 2849-2856.

6. Lee DY, An G (2012) Two AP2 family genes, supernumerary bract (SNB) and Osindeterminate spikelet 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. Plant J 69: 445-461.

7. Vollbrecht E, Springer PS, Goh L, Buckler ESt, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. Nature 436: 1119-1126.

8. Davidson RM, Hansey CN, Gowda M, Childs KL, Lin H, et al. (2011) Utility of RNA sequencing for analysis of maize reproductive transcriptomes. The Plant Genome 4(3): 191-203.

9. Zhuang Y, Ren G, Yue G, Li Z, Qu X, et al. (2007) Effects of water-deficit stress on the transcriptomes of developing immature ear and tassel in maize. Plant Cell Rep 26: 2137-2147.

10. Liu H, Qin C, Chen Z, Zuo T, Yang X, et al. (2014) Identification of miRNAs and their target genes in developing maize ears by combined small RNA and degradome sequencing. BMC Genomics 15: 25.

11. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326: 1112-1115.

12. Eveland AL, Satoh-Nagasawa N, Goldshmidt A, Meyer S, Beatty M, et al. (2010) Digital gene expression signatures for maize development. Plant Physiol 154: 1024-1039.

13. Veitch NJ, Johnson PC, Trivedi U, Terry S, Wildridge D, et al. (2010) Digital gene expression analysis of two life cycle stages of the human-infective parasite, Trypanosoma brucei gambiense reveals differentially expressed clusters of co-regulated genes. BMC Genomics 11: 124.

14. Tao X, Gu YH, Wang HY, Zheng W, Li X, et al. (2012) Digital gene expression analysis based on integrated de novo transcriptome assembly of sweet potato [Ipomoea batatas (L.) Lam]. PLoS One 7: e36234.

15. Zhang X, Hao L, Meng L, Liu M, Zhao L, et al. (2013) Digital gene expression tag profiling analysis of the gene expression patterns regulating the early stage of mouse spermatogenesis. PLoS One 8: e58680.

16. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63.

17. Kakumanu A, Ambavaram MM, Klumas C, Krishnan A, Batlang U, et al. (2012) Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-Seq. Plant Physiol 160: 846-867.

18. Sekhon RS, Briskine R, Hirsch CN, Myers CL, Springer NM, et al. (2013) Maize gene atlas developed by RNA sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. PLoS One 8: e61005.

19. Li H, Peng Z, Yang X, Wang W, Fu J, et al. (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45: 43-50.

20. Shen Y, Jiang Z, Yao X, Zhang Z, Lin H, et al. (2012) Genome expression profile analysis of the immature maize embryo during dedifferentiation. PLoS One 7: e32237.

21. Zhao Y, Xu Z, Mo Q, Zou C, Li W, et al. (2013) Combined small RNA and degradome sequencing reveals novel miRNAs and their targets in response to low nitrate availability in maize. Ann Bot 112: 633-642.

22. Shen Y, Zhang Y, Chen J, Lin H, Zhao M, et al. (2013) Genome expression profile analysis reveals important transcripts in maize roots responding to the stress of heavy metal Pb. Physiol Plant 147: 270-282.

23. Zhu Y, Fu J, Zhang J, Liu T, Jia Z, et al. (2009) Genome-wide analysis of gene expression profiles during ear development of maize. Plant Mol Biol 70: 63-77.

24. Li R, Yu C, Li Y, Lam TW, Yiu SM, et al. (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25: 1966-1967.

25. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125: 279-284.

26. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37: 914-939.

27. Krishna K, Narasimha Murty M (1999) Genetic K-means algorithm. IEEE Trans Syst Man Cybern B Cybern 29: 433-439.

28. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

29. Vollbrecht E, Schmidt R (2009) Development of the Inflorescences. Handbook of Maize: Its Biology New York: Springer. pp. 13-40.

30. Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, et al. (2009) EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 19: 327-335.

31. Hou H, Li J, Gao M, Singer SD, Wang H, et al. (2013) Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. PLoS One 8: e59358.

32. Salinas M, Xing S, Hohmann S, Berndtgen R, Huijser P (2012) Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. Planta 235: 1171-1184.

33. Ma J, Morrow DJ, Fernandes J, Walbot V (2006) Comparative profiling of the sense and antisense transcriptome of maize lines. Genome Biol 7: R22.

34. Drews GN, Bowman JL, Meyerowitz EM (1991) Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. Cell 65: 991-1002.

35. Sridhar VV, Surendrarao A, Gonzalez D, Conlan RS, Liu Z (2004) Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. Proc Natl Acad Sci U S A 101: 11494-11499.

36. Dykema PE, Sipes PR, Marie A, Biermann BJ, Crowell DN, et al. (1999) A new class of proteins capable of binding transition metals. Plant Mol Biol 41: 139-150.

37. Hernandez-Domiguez EE, Valencia-Turcotte LG, Rodriguez-Sotres R (2012) Changes in expression of soluble inorganic pyrophosphatases of Phaseolus vulgaris under phosphate starvation. Plant Sci 187: 39-48.

38. Ruszkowski M, Brzezinski K, Jedrzejczak R, Dauter M, Dauter Z, et al. (2013) Medicago truncatula histidine-containing phosphotransfer protein: structural and biochemical insights into the cytokinin transduction pathway in plants. FEBS J 280: 3709-3720.

39. Quilis J, Lopez-Garcia B, Meynard D, Guiderdoni E, San Segundo B (2013) Inducible expression of a fusion gene encoding two proteinase inhibitors leads to insect and pathogen resistance in transgenic rice. Plant Biotechnol J.

40. King GJ, Chanson AH, McCallum EJ, Ohme-Takagi M, Byriel K, et al. (2013) The Arabidopsis B3 domain protein VERNALIZATION1 (VRN1) is involved in processes essential for development, with structural and mutational studies revealing its DNA-binding surface. J Biol Chem 288: 3198-3207.

Figure Legends and Tables

Figure 1. Comparison of four development stages of maize ear. Comparison of genes expressed in sense (A) and antisense (B) directions in the four development stages. Overlaps show the number of genes shared between stages. (C) A Venn diagram shows the genes expressed in sense and antisense direction in each developmental stage, and all stages.

Figure 2. Genes differentially expressed during various comparisons between developmental stages. (A) Data were shown as percentages and numbers of up- and down-regulated genes within each comparison. (B) Venn diagram of differentially expressed genes with adjacent comparisons.

Figure 3. Results of gene expression validated by quantitative real-time PCR analysis. Genes were selected based on sequencing results.

Figure 4. Dynamic progression of maize ear transcriptome. (A) Dynamic progression of ear transcriptome. 18 clusters were identified along the four developmental stages from 6,800 differentially expressed genes. The 18 clusters are presented in A. (B) Functional category enrichment among the 18 clusters. (C) Pathways enrichment analyses among the 18 clusters.

Figure 5. miR156 and its predicted targets SBP-box genes in maize ear development (Adapted from Liu *et al.* (2014)). (A) The development of the maize ear regulated by miR156 and its predicted targets SBP-box genes, *SBP-5* and *SBP-6*. Both of *SBP-5* and *SBP-6* were found to be targets of miR156, during all four stages of maize ear development [10]. (B) The expression pattern of *SBP-5* and *SBP-6*. During ear development. *SBP-5* was significantly up-regulated in stage II (spikelet differentiation phase) and III (floret primordium differentiation phase), while *SBP-6* was up-regulated instead of decreased expression of *SBP-5* in stage IV (floret organ differentiation). (C) The expression pattern of *SBP* (GRMZM2G109354). As a transcriptional factor, SBP

(GRMZM2G109354) is also found to be significantly up- or down-regulated during ear development, suggesting a potential regulatory role of SBP in the maize ear development. * indicated the significant difference in adjacent two developmental stages ($p < 0.01$, $n = 3$).

Figure S1. Distribution of clean tags mapped to the reference genome during four developmental stages.

Figure S2. Differentially expressed genes of each comparison among the four developmental stages.

Figure S3. Comparison of gene expression between qRT-PCR and Solexa sequencing analyses. Heatmap and Pearson correlation coefficients of 32 selected genes between these two analyses were shown in (A) and (B). "PCR-geneID" and "TPM-geneID" represented gene expression analyzed by qRT-PCR and Solexa sequencing, respectively.

Figure S4. Selected genes related to maize ear development. "#" represents gene annotation or *mutant* (Italic). "**" represents false discovery rate < 0.001 with comparisons between the adjacent developmental stages.

Table S1. Reference tag database and major characteristics of DGE libraries.

Table S2. Gene expression in both antisense and sense direction during developmental stages of maize ear.

Table S3. Expression of all genes across four developmental stages.

Table S4. Differentially expressed genes between different libraries.

Table S5. Functional category analysis and pathway enrichment analysis of differentially expressed genes within 18 clusters.

Table S6. Transcriptional factors expressed during the four developmental stages.

Table S7. Differentially expressed transcriptional factors between stages during the four developmental stages of maize ear.

Table S8. Primers of real-time RT-PCR assay used in this study.xlsx

Tables

Table 1. Statistics of tags mapping to maize genome during four stages

Summary	Stage I	Stage II	Stage III	Stage IV
Raw Data	4200000	4200000	4200000	4200000
Distinct Raw Data*	290638	304701	286788	267684
Clean Tag	3931920	4066158	4054326	4055577
Distinct Clean Tag	270298	278664	256999	238767
Clean Tag/Raw Data	93.62%	96.81%	96.54%	96.56%
All Tag Mapping to Gene	2662565	2678767	2920344	2991574
All Tag Mapping to Gene*	67.72%	65.88%	72.03%	73.76%
All Tag Mapping to Gene	130940	137946	138422	133110
All Tag Mapping to Gene*	48.44%	49.50%	53.86%	55.75%
Unambiguous Tag Mapping to Gene	2388758	2379223	2646049	2690916
Unambiguous Tag Mapping to Gene*	60.75%	58.51%	65.26%	66.35%
Unambiguous Tag Mapping to Gene	115681	122033	122663	117943
Unambiguous Tag Mapping to Gene*	42.80%	43.79%	47.73%	49.40%
All Tag-mapped Genes	24814	24999	24754	24480
All Tag-mapped Genes**	76.26%	76.83%	76.07%	75.23%
Unambiguous Tag-mapped Genes	21559	21669	21525	21257
Unambiguous Tag-mapped Genes**	66.25%	66.59%	66.15%	65.33%
Unknown Tag	789178	891010	647233	591946
Unknown Tag*	20.07%	21.91%	15.96%	14.60%
Unknown Tag	96886	98045	78486	68094
Unknown Tag*	35.84%	35.18%	30.54%	28.52%

*% of Clean Tag; **% of Ref genes;
Statistics of raw tags, clean tags, tags mapped to genes, unambiguous tags and unknown tags.

Table 2 Predicted differentially expressed genes targeted by microRNA found by Liu *et al.* (2014)

GeneID	microRNA[a]	log2FC[b]	q value	Gene Description
stage II vs. stage I				
GRMZM2G067624	miR156	1.97	0.00E+00	unknown [*Zea mays*]
GRMZM2G163813	miR156	-2.95	4.12E-38	unknown [*Zea mays*]
GRMZM2G160917	miR156	1.85	2.02E-8	SBP-domain protein 5 [*Zea mays*]
GRMZM2G089361	miR319	1.47	8.62E-10	hypothetical protein SORBIDRAFT_01g006020 [*Sorghum bicolor*]
stage III vs. stage II				
GRMZM2G126018	miR156	1.02	2.67E-04	unknown [*Zea mays*]
GRMZM2G160917	miR156	1.72	0.00E+00	SBP-domain protein 5 [*Zea mays*]
GRMZM2G163813	miR156	1.49	1.89E-06	unknown [*Zea mays*]
GRMZM2G081406	miR160	3.52	8.60E-05	hypothetical protein SORBIDRAFT_04g026610 [*Sorghum bicolor*]
GRMZM2G153233	miR160	1.71	6.09E-07	hypothetical protein LOC100304210 [*Zea mays*]
GRMZM2G393433	miR164	2.59	6.36E-04	hypothetical protein [*Zea mays*]
GRMZM2G028980	miR167	1.43	1.57E-06	hypothetical protein SORBIDRAFT_04g004430 [*Sorghum bicolor*]
GRMZM2G155490	miR390	3.44	2.22E-04	-
GRMZM2G443903	miR396	3.21	1.40E-04	putative pol protein [*Zea mays*]
stage VI vs. stage III				
GRMZM2G307588	miR156	1.37	1.52E-09	SBP-domain protein 6 [*Zea mays*]
GRMZM2G460544	miR156	-2.54	3.79E-51	unknown [*Zea mays*]
GRMZM2G081406	miR160	1.13	3.18E-07	hypothetical protein SORBIDRAFT_04g026610 [*Sorghum bicolor*]
GRMZM2G089361	miR319	-2.07	7.09E-10	hypothetical protein SORBIDRAFT_01g006020 [*Sorghum bicolor*]
GRMZM2G109843	miR319	-1.23	8.85E-10	hypothetical protein [*Zea mays*]
GRMZM2G131280	miR529	-1.36	2.06E-05	hypothetical protein LOC100277728 [*Zea mays*]
GRMZM2G136158	miR529	1.58	8.41E-05	hypothetical protein SORBIDRAFT_03g010740 [*Sorghum bicolor*]

[a] microRNAs identified by Liu H *et al.* (2014) [10].

[b] "FC" represents "fold change".

Table 3 Differentially expressed TFs across three comparisons of adjacent developmental stages

Gene	PlantTFDB_ID	TF_Family	IIvs.I[a]	IIIvs.II[b]	IVvs.III[c]	Gene Description
GRMZM2G065496	Zma029754	B3	-1.08	1.51	-1.00	B3 DNA binding domain containing protein [*Zea mays*]
GRMZM2G052102	Zma030311	bZIP	-1.40	-1.66	-3.92	hypothetical protein SORBIDRAFT_04g008840 [*Sorghum bicolor*]
GRMZM2G479885	Zma028890	bZIP	-1.08	1.73	1.05	unknown [*Zea mays*]
GRMZM2G061487	Zma003671	ERF	-2.79	2.68	-1.45	DRE binding factor 1 [*Zea mays*]
GRMZM2G056600	Zma017283	HD-ZIP	1.56	-1.20	2.74	hypothetical protein LOC100272620 [*Zea mays*]
GRMZM2G044902	Zma026448	LBD	-1.43	1.23	-1.19	hypothetical protein SORBIDRAFT_01g031790 [*Sorghum bicolor*]
GRMZM2G129034	Zma050175	MIKC	-2.20	2.90	2.38	unknown [*Zea mays*]
GRMZM2G137510	Zma056196	MIKC	-1.21	-1.38	-4.24	unknown [*Zea mays*]
GRMZM2G050550	Zma002240	MYB	-1.95	3.36	-1.27	sucrose responsive element binding protein [*Zea mays*]
GRMZM2G127379	Zma007036	NAC	-1.34	1.86	-2.49	unknown [*Zea mays*]
GRMZM2G347043	Zma057817	NAC	-1.88	-1.26	-1.81	NAC1 transcription factor [*Zea mays*]
GRMZM2G109354	Zma006127	SBP	-1.24	1.32	-1.25	H0215A08.3 [*Oryza sativa (indica cultivar-group)*]
GRMZM2G113888	Zma001368	TCP	-2.05	3.27	1.22	hypothetical protein LOC100272799 [*Zea mays*]

[a,b,c] The values indicate the log2 fold change